

# TWO-COLOR BALANCED AFFINE URN MODELS WITH MULTIPLE DRAWINGS II: LARGE-INDEX AND TRIANGULAR URNS

MARKUS KUBA AND HOSAM M. MAHMOUD

**ABSTRACT.** This is the second part of a two-part investigation. We continue the study of a class of balanced urn schemes on balls of two colors (white and black). At each drawing, a sample of size  $m \geq 1$  is drawn from the urn and ball addition rules are applied; the special case  $m = 1$  of sampling only a single ball coincides with ordinary balanced urn models. We consider these multiple drawings under sampling with or without replacement. For the class of affine conditional expected value, we study the number of white balls after  $n$  steps. The affine class is parametrized by  $\Lambda$ , specified by the ratio of the two eigenvalues of a reduced ball replacement matrix and the sample size, leading to three different cases: small-index urns ( $\Lambda \leq \frac{1}{2}$ , and the case  $\Lambda = \frac{1}{2}$  is critical), large-index urns ( $\Lambda > \frac{1}{2}$ ), and triangular urns. In Part I we derived a central limit theorem for small index urns, and proved almost-sure convergence for large index and triangular urn models. In the present paper (Part II), we continue the study of affine urn schemes and study the moments of large-index urns and triangular urn models. We show moment convergence under suitable scaling and we also provide expressions for the moments.

## 1. INTRODUCTION

Urn schemes are simple, yet versatile mathematical tools for modeling evolutionary processes. This is the second part of a two-part investigation of affine urn schemes. We discussed motivation and application in the first part; we refer the interested reader to the introduction of that companion paper. This investigation is devoted to the study of a generalization of the two-color Pólya urn model, where *multiple* balls are drawn at each discrete time step, their colors are inspected, then the entire sample is placed back in the urn, and rules of replacement are affected. Such urn schemes with multiple drawing recently received attention [5, 6, 16, 19, 23, 25, 27, 29]. The addition/removal of balls depends on the combinations of colors in the multiset drawn. We suppose the two colors are white (W) and black (B), and use the notation  $\{W^k B^{m-k}\}$  to refer to a sample of size  $m$  containing  $k$  white balls and  $m - k$  black balls. Specifically, we draw  $m \geq 1$  balls and add/remove white and black balls according to the multiset  $\{W^k B^{m-k}\}$  of observed colors: If we draw  $k$  white and  $m - k$  black balls, we add  $a_{m-k}$  white and  $b_{m-k}$  black balls,  $0 \leq k \leq m$ . The ball replacement matrix of this urn model with multiple drawings is a rectangular  $(m+1) \times 2$

---

*Date:* October 1, 2015.

*2000 Mathematics Subject Classification.* 60C05, 60F05, 60G42.

*Key words and phrases.* Urn model, random structure, martingale, variance, central limit theorem, large-index urns, triangular urns.

matrix:

$$\mathbf{M} = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \\ \vdots & \vdots \\ a_{m-1} & b_{m-1} \\ a_m & b_m \end{pmatrix}. \quad (1)$$

Let  $W_n$  and  $B_n$  be respectively the number of white and black urns in the urn after  $n$  draws. We assume throughout that the urn model is *balanced*, a scheme in which one adds a constant number of balls  $\sigma \geq 1$  each time, regardless of what multiset is withdrawn. That is,  $a_k + b_k = \sigma \geq 1$ ,  $0 \leq k \leq m$ . Consequently, the total number of balls  $T_n = W_n + B_n$  after  $n$  draws is given by a nonrandom number  $T_n = T_0 + \sigma n$ . We confine our attention to the so-called tenable urn models, where the process of drawing and replacing balls can be continued ad infinitum.

Urn models with multiple drawings and sample size  $m \geq 2$  are in general more difficult to analyze compared to the ordinary case  $m = 1$ , see [5, 6, 23, 25, 27] and the discussions therein. In Part I we carried out a structural analysis classifying urn models with multiple drawings according to the shape of the conditional expected value of the number of white balls  $W_n$ . In this work, as in Part I, we analyze urn schemes with multiple drawings for which the conditional expectation of the number of white balls  $W_n$  after  $n$  draws has an affine structure of the form

$$\mathbb{E}[W_n | \mathbb{F}_{n-1}] = \alpha_n W_{n-1} + \beta_n, \quad n \geq 1. \quad (2)$$

Here,  $\alpha_n, \beta_n$  denote deterministic sequences certain sequences depending only on  $n$ ,  $a_{m-1}$ ,  $a_m$ , and the balance factor  $\sigma$ , and  $\mathbb{F}_n$  denotes the sigma-algebra generated by the first  $n$  draws from the urn. Note that in the case  $m = 1$  balanced urn models are by definition affine.

**1.1. Affine schemes.** In Part I we obtained a characterization of affine schemes. Basically, a scheme is *affine*, if and only if the entries in the first column satisfy the recurrence:

$$a_k = (m - k)(a_{m-1} - a_m) + a_m, \quad \text{for } 0 \leq k \leq m. \quad (3)$$

Equivalently, the condition can be written as

$$a_k = hk + a_0,$$

for arbitrary  $h$  respecting tenability. The values  $\alpha_n$  and  $\beta_n$  in (2) are given by

$$\alpha_n = \frac{T_{n-1} + m(a_{m-1} - a_m)}{T_{n-1}}, \quad \beta_n = a_m, \quad n \geq 1. \quad (4)$$

In view of the balance, the entries in the second column satisfy a similar recurrence.

**1.2. Classification of limit laws.** We also discussed in Part I a classification in terms of the two eigenvalues  $\Lambda_1$  and  $\Lambda_2$  of the submatrix  $\begin{pmatrix} a_{m-1} & b_{m-1} \\ a_m & b_m \end{pmatrix}$  with  $\Lambda_1$  being the larger of the two eigenvalues.<sup>1</sup> The urn index is

$$\Lambda = \Lambda_2/\Lambda_1 = \frac{m}{\sigma}(a_{m-1} - a_m).$$

A trichotomy of cases similar to the classical case of sample size  $m = 1$  arises: (1) Urn schemes with  $a_m \neq 0$  and a small index. These have  $\Lambda \leq \frac{1}{2}$ , and the case  $\Lambda = \frac{1}{2}$  is critical, (2) Urn schemes with  $a_m \neq 0$  and a large index,  $\Lambda > \frac{1}{2}$ , (3) Triangular urn models with  $a_m = 0$ . The limiting distribution results for sample size  $m = 1$  and balanced urn models are well known and are, amongst others, based on the results of Bagci and Pal [2], Janson [13, 14], Flajolet et al. [9], Chauvin et al. [3] and Neininger and Knappe [17]. For a balanced two-color urn model  $M = \begin{pmatrix} a_0 & b_0 \\ a_1 & b_1 \end{pmatrix}$  let  $\Lambda$  denote the ratio of the two eigenvalues of  $M$ . For small-index urns,  $\Lambda \leq \frac{1}{2}$ ,

one obtains a central limit theorem for the number of white balls  $\frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathbb{V}(W_n)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ . For large-index urns,  $\Lambda > \frac{1}{2}$ , and also triangular urns,  $b_0 \cdot a_1 = 0$ , the suitably normalized (and centered) number of white balls converges almost surely to a non-normal limiting distribution. Concerning the distribution of the limit law of large index urns there has been a *flurry of activity* in the last decade. Several articles [3, 4, 26] have been entirely devoted to the study of large-index urns and the properties of the limit law such as characteristic function, moments, decomposition of random variables, fixed-point equations (smoothing transforms) etc.; see also [13, 17] for general studies including large-index urn models; moreover, triangular urn models and properties of its limit law, including moments and density functions, have been analyzed in [9, 14].

We will extend the classification result above for  $m = 1$  to arbitrary affine urn models. In Part I we already discussed the case  $\Lambda \leq \frac{1}{2}$  and proved a central limit theorem. Moreover, for  $\Lambda > \frac{1}{2}$  and for triangular urns almost-sure convergence of suitably defined random variables have been proved in part I using discrete martingales. We continue our investigation analyzing the positive integer moments of the limit laws of large-index urns and triangular urns. It turns out that the higher moments of the limiting distributions of large-index urn models and triangular urn models are given by nested infinite sums. Our results are valid for any  $m \geq 1$ , but in contrast to the case  $m = 1$  it appears that in general for  $m \geq 2$  the nested sums cannot be simplified.

Note that concerning urn models with multiple drawings and replacement matrix  $M$  as given by (1) we call an urn model *triangular* if  $a_m = 0$  or  $b_0 = 0$  or both  $a_m = b_0 = 0$ . The case  $b_0 = 0$  and  $a_m > 0$  for the black balls corresponds to the case  $a_m = 0$  and  $b_0 > 0$  for the white balls; by the relation  $B_n = T_n - W_n$  this implies that, without loss of generality, we can restrict our attention to triangular urns with  $a_m = 0$  and  $b_0 \geq 0$ . If both  $a_m = b_0 = 0$  one obtains the so-called Pólya urn model as treated in [6]. Setting  $a_{m-1} = c > 0$  triangular urns in the affine

---

<sup>1</sup>An equivalent formulation can be obtained from the first two rows of the replacement matrix; see part I and for a discussion of the higher dimensional condition.

scheme with  $a_m = 0$  are specified by a rectangular  $(m + 1) \times 2$  matrix:

$$\mathbf{M} = \begin{pmatrix} mc & \sigma - mc \\ (m-1)c & \sigma - (m-1)c \\ \dots & \dots \\ c & \sigma - c \\ 0 & \sigma \end{pmatrix},$$

and the three parameters  $a_{m-1} = c > 0$ , the sample size  $m$  and the total balance  $\sigma > 0$  such that  $\sigma \geq mc$ . The special case  $\sigma = mc$  corresponds to generalized Pólya urn model, as discussed in [6]. We will improve the results of [6] by obtaining an explicit nonrecursive formula for the moments of  $W_n$ , as well as for the limiting distribution.

**1.3. Affine urn models.** A direct consequence of the affine expectation (2) is a martingale structure, and a closed form expression for the expected value and the second moment. We collect the results relevant for this work in the following Lemma.

**Lemma 1** ([20]). *The expected value of the random variable  $W_n$ , counting the number of white balls in a two-color balanced affine urn model with multiple drawings, is for both sampling models  $\mathcal{M}$  and  $\mathcal{R}$  given by  $\mathbb{E}[W_n] = \frac{a_m}{g_n} \sum_{j=1}^n g_j + W_0 \frac{1}{g_n}$ . Here, we have*

$$g_n = \prod_{j=0}^{n-1} \frac{T_j}{T_j + m(a_{m-1} - a_m)} = \frac{\binom{n-1+\frac{T_0}{\sigma}}{n}}{\binom{n-1+\frac{T_0}{\sigma}+\Lambda}} = \frac{\Gamma(\frac{T_0}{\sigma} + \Lambda)}{\Gamma(\frac{T_0}{\sigma})} n^{-\Lambda} \left(1 + O\left(\frac{1}{n}\right)\right). \quad (5)$$

For  $a_m \neq 0$  and  $\Lambda < 1$ , including large-index urns, we have

$$\begin{aligned} \mathbb{E}[W_n] &= \frac{a_m(n + \frac{T_0}{\sigma})}{1 - \Lambda} + \left(W_0 - \frac{a_m T_0}{1 - \Lambda}\right) \frac{\binom{n-1+\frac{T_0}{\sigma}+\Lambda}{n}}{\binom{n-1+\frac{T_0}{\sigma}}{n}} \\ &= \frac{a_m}{1 - \Lambda} n + \left(W_0 - \frac{a_m T_0}{1 - \Lambda}\right) \frac{\Gamma(\frac{T_0}{\sigma})}{\Gamma(\frac{T_0}{\sigma} + \Lambda)} n^{\Lambda} + \mathcal{O}(1). \end{aligned}$$

For triangular urns with  $a_m = 0$  we have the closed form expression

$$\mathbb{E}[W_n] = W_0 \frac{n\sigma + T_0}{T_0}.$$

The random variable  $\mathcal{W}_n = g_n(W_n - \mathbb{E}[W_n])$  is a centered martingale with respect to the natural filtration:  $\mathbb{E}[\mathcal{W}_n | \mathbb{F}_{n-1}] = \mathcal{W}_{n-1}$ ,  $n \geq 1$ , with  $\mathcal{W}_0 = 0$ . For large-index urns,  $\mathcal{W}_n$  converges almost surely and in  $L_2$  to a limit  $\mathcal{W}_{\infty}$ . For triangular urn models, where  $a_m = 0$ , the random variable  $\mathfrak{W}_n = g_n W_n$  is a nonnegative martingale and converges almost surely to a limit  $\mathfrak{W}_{\infty}$ .

**Remark 1.** A slight unpleasant complication is the case distinction for the expected value between  $a_m \neq 0$  and  $a_m = 0$ . The case  $a_m = 0$  leads to  $\Lambda = \frac{ma_{m-1}}{\sigma}$ , which may be equal to one for Pólya urn models with  $a_{m-1} = c$  and  $\sigma = mc$ . However, we can interpret the explicit formula  $\mathbb{E}[W_n]$  stated for  $a_m \neq 0$  and  $\Lambda < 1$  in the right way: We set first  $a_m = 0$ , so all terms which

include  $\frac{1}{1-\Lambda}$  vanish, and only afterward we set  $\Lambda$  to its corresponding value, including the case  $\Lambda = 1$ . In other words, the quotient  $\frac{a_m}{1-\Lambda}$  is zero for  $a_m = 0$ , regardless of the value of  $\Lambda \leq 1$ .

**1.4. Plan of the paper and notation.** In Part I we obtained Gaussian limits for small-index urn schemes and almost-sure limits for triangular and large-index urns. It is our aim to complete the study of triangular and large-index urns. We study the (positive integer) moments  $W_n^s$ ,  $s \in \mathbb{N}$ , of  $W_n$  and the moments of the limit laws  $\mathcal{W}_\infty$  for large-index urns and  $\mathfrak{W}_\infty$  for triangular urns. We provide a recursive characterization for the moments of the limiting distribution. For triangular urns  $a_m = 0$  and  $b_0 \geq 0$ , we generalize the existing results concerning the case  $a_m = b_0 = 0$  (see [6]) obtaining an explicit nonrecursive descriptions of the moments.

We denote by  $x^{\underline{k}}$  the  $k$ th falling factorial,  $x(x-1)\dots(x-k+1)$ ,  $k \geq 0$ , with  $x^0 = 1$ . We shall also use  $\nabla$ , the backward difference operator, defined by  $\nabla h_n = h_n - h_{n-1}$ , when acting on a function  $h_n$ . We use  $\left\{ \begin{smallmatrix} s \\ k \end{smallmatrix} \right\}$  to denote the Stirling numbers of the second kind, and  $\left[ \begin{smallmatrix} s \\ k \end{smallmatrix} \right]$  to denote the unsigned Stirling numbers of the first kind (see [28] or [11]); these numbers appear as coefficients in the expansions

$$x^s = \sum_{k=0}^s \left\{ \begin{smallmatrix} s \\ k \end{smallmatrix} \right\} x^{\underline{k}}, \quad x^{\underline{s}} = \sum_{k=0}^s (-1)^{s-k} \left[ \begin{smallmatrix} s \\ k \end{smallmatrix} \right] x^k,$$

relating ordinary powers  $x^s$  to the falling factorials  $x^{\underline{s}}$ . Moreover, in this article we refer with  $\mathcal{W}_\infty$  to the almost-sure limit of  $W_n$  for large-index urns with  $\frac{1}{2} < \Lambda < 1$  and with  $\mathfrak{W}_\infty$  to the almost-sure limit for triangular urns satisfying  $a_m = 0$  and  $b_0 \geq 0$ .

## 2. PRELIMINARIES

**2.1. Sampling schemes.** Assume that an urn contains  $w$  white and  $b$  black balls. We consider two different sampling schemes for drawing the  $m$  balls at each step: model  $\mathcal{M}$  and model  $\mathcal{R}$ . In model  $\mathcal{M}$ , we draw the  $m$  balls without replacement. The  $m$  balls are drawn at once and their colors are examined. After the sample is collected, we put the entire sample back in the urn and execute the replacement rules according to the counts of colors observed. The tenability assumption implies that for model  $\mathcal{M}$  the coefficients  $a_k$  of the ball replacement matrix (1) satisfy the condition  $a_k \geq -(m-k)$ ,<sup>2</sup> for  $0 \leq k \leq m$ . Without loss of generality we assume throughout this work that the initial number of balls  $T_0 = W_0 + B_0 \geq m$ .

The probability  $\mathbb{P}(W^k B^{m-k})$  of drawing  $k$  white and  $m-k$  black balls is given by

$$\mathbb{P}(W^k B^{m-k}) = \frac{1}{(b+w)^{\underline{m}}} \binom{m}{k} w^{\underline{k}} b^{\underline{m-k}} = \frac{\binom{w}{k} \binom{b}{m-k}}{\binom{b+w}{m}}, \quad 0 \leq k \leq m.$$

Thus  $X$ , the number of white balls in the sample, follows a hypergeometric distribution, with parameters  $w+b$ ,  $w$ , and  $m$ , that is, one that counts the number of white balls in a sample of

<sup>2</sup>These assumptions can be relaxed a little bit, if the initial values  $W_0$  and  $B_0$  are adapted to the entries in the ball replacement matrix. E.g., for  $m=1$  the urn model with ball replacement matrix  $\begin{pmatrix} -3 & 8 \\ 6 & -4 \end{pmatrix}$  is still tenable if  $W_0$  is a multiple of 3 and  $B_0$  a multiple of 4.

size  $m$  balls taken out of an urn containing  $w$  white and  $b$  black balls (a total of  $\tau = w + b$  balls). The first two moments of  $X$  are given by

$$\mathbb{E}[X] = m \frac{w}{\tau}, \quad \mathbb{E}[X^2] = \frac{w(w-1)m(m-1)}{\tau(\tau-1)} + \frac{wm}{\tau}.$$

The  $\ell$ th moments of  $X$  can be written as a polynomial in the parameter  $w$ :

$$\mathbb{E}[X^\ell] = \sum_{k=0}^m k^\ell \frac{\binom{w}{k} \binom{\tau-w}{m-k}}{\binom{\tau}{m}} = \sum_{i=0}^{\ell^*} w^i \sum_{j=i}^{\ell^*} (-1)^{j-i} \frac{\begin{bmatrix} j \\ i \end{bmatrix} \left\{ \begin{smallmatrix} \ell \\ j \end{smallmatrix} \right\} \binom{m}{j}}{\binom{\tau}{j}},$$

with  $\ell^* = \min\{\ell, m\}$ , for arbitrary  $\ell \geq 0$ .

In model  $\mathcal{R}$ , we draw the  $m$  balls with replacement. The  $m$  balls are drawn one at a time. After a ball is drawn, its color is observed, and is reinserted in the urn, and thus it might reappear in the sampling of one multiset. After  $m$  balls are collected in this way (and they are all back in the urn), we execute the replacement rules according to the counts of colors observed. By the tenability assumption  $a_k \geq -1$  for  $0 \leq k \leq m-1$  and  $a_m \geq 0$  for model  $\mathcal{R}$ .

The probability  $\mathbb{P}(W^k B^{m-k})$  of drawing  $k$  white and  $m-k$  black balls is given by

$$\mathbb{P}(W^k B^{m-k}) = \frac{1}{(b+w)^m} \binom{m}{k} w^k b^{m-k}, \quad 0 \leq k \leq m.$$

In other words, under model  $\mathcal{R}$ , the number of white balls in the multiset of size  $m$  follows a binomial distribution with parameters  $m$ , and  $w/\tau$ , one that counts the number of successes in  $m$  independent identically distributed experiments, with  $w/\tau$  probability of success per experiment. Let  $Y$  denote such a binomially distributed random variable. Then, the first two moments of  $Y$  are given by

$$\mathbb{E}[Y] = m \frac{w}{\tau}, \quad \mathbb{E}[Y^2] = m \frac{w}{\tau} \left(1 - \frac{w}{\tau}\right) + m^2 \frac{w^2}{\tau^2}.$$

The  $\ell$ th moments of  $Y$  can be written as a polynomial in the parameter  $w$ :

$$\mathbb{E}[Y^\ell] = \sum_{k=0}^m k^\ell \binom{m}{k} \frac{w^k (\tau-w)^{m-k}}{\tau^m} = \sum_{j=0}^{\ell} \left\{ \begin{smallmatrix} \ell \\ j \end{smallmatrix} \right\} m^j \frac{w^j}{\tau^j}.$$

**2.2. Distributional equations.** In what follows, we use the notation  $\mathbb{I}_n(W^k B^{m-k})$  to stand for the indicator of the event that the multiset  $\{W^k B^{m-k}\}$  is drawn in the  $n$ th sampling. Conditioning on the outcome of the  $n$ th draw, we obtain a distributional equation for  $W_n$ . The number of white balls after  $n$  draws is the number of white balls after  $n-1$  draws, plus the contribution of white balls after the  $n$ th sample is obtained (with  $\stackrel{\mathcal{L}}{=}$  for equality in law):

$$W_n \stackrel{\mathcal{L}}{=} W_{n-1} + \sum_{k=0}^m a_{m-k} \mathbb{I}_n(W^k B^{m-k}), \quad n \geq 1. \quad (6)$$

Let  $\mathbb{F}_{n-1}$  denote the  $\sigma$ -field generated by the first  $n - 1$  draws. The indicators  $\mathbb{I}_n(W^k B^{m-k})$  satisfy

$$\mathbb{P}(\mathbb{I}_n(W^k B^{m-k}) = 1 \mid \mathbb{F}_{n-1}) = \frac{\binom{W_{n-1}}{k} \binom{B_{n-1}}{m-k}}{\binom{T_{n-1}}{m}} = \frac{\binom{W_{n-1}}{k} \binom{T_{n-1}-W_{n-1}}{m-k}}{\binom{T_{n-1}}{m}} \quad (7)$$

for model  $\mathcal{M}$ , and

$$\mathbb{P}(\mathbb{I}_n(W^k B^{m-k}) = 1 \mid \mathbb{F}_{n-1}) = \binom{m}{k} \frac{W_{n-1}^k B_{n-1}^{m-k}}{T_{n-1}^m} = \binom{m}{k} \frac{W_{n-1}^k (T_{n-1} - W_{n-1})^{m-k}}{T_{n-1}^m} \quad (8)$$

for model  $\mathcal{R}$ . We obtain for  $W_n^s$ ,  $s \geq 1$ , a distributional equation by taking the  $s$ th power of (6), and using the fact that the indicator variables are mutually exclusive:

$$W_n^s \stackrel{\mathcal{L}}{=} \sum_{\ell=0}^s \binom{s}{\ell} W_{n-1}^{s-\ell} \sum_{k=0}^m a_{m-k}^\ell \mathbb{I}_n(W^k B^{m-k}), \quad n \geq 1. \quad (9)$$

### 3. MOMENT STRUCTURE

In order to study the moments of  $W_n$  and of the almost sure limits  $\mathcal{W}_\infty$ ,  $\mathfrak{W}_\infty$ , we analyze higher shifted moments. This will enable us to provide a recursive characterization of the moments of both. We also complement these result by obtaining concrete explicit expressions for the moments.

**3.1. Higher moments: exact representations.** The asymptotic expansion of the expected value suggests that we shift  $W_n$  by the dominant term  $\frac{a_m}{1-\Lambda} n$  of its asymptotic expansion (1). Thus, we consider the shifted random variable  $\tilde{W}_n$  defined by

$$\tilde{W}_n = W_n - \frac{a_m}{1-\Lambda} n, \quad (10)$$

which is well defined for arbitrary  $\Lambda < 1$ . For triangular urn models  $a_m = 0$ , so

$$\tilde{W}_n = W_n. \quad (11)$$

As discussed in Remark 1, this random variable has to be interpreted in the right way in the Pólya urn case when  $a_m = 0$  and  $\Lambda = 1$ , interpreting the fraction  $\frac{a_m}{1-\Lambda}$  as zero, so  $\tilde{W}_n = W_n$ , regardless of the value of  $\Lambda$ . We obtain the following recurrence relation for the moments  $\mathbb{E}[\tilde{W}_n^s]$ .

**Lemma 2.** *The positive integer moments  $\mathbb{E}[\tilde{W}_n^s]$  of the shifted random variable  $\tilde{W}_n = W_n - \frac{a_m}{1-\Lambda} n$ , with  $W_n$  counting the number of white balls in a two-color balanced affine urn model with multiple drawings, satisfy for both sampling models  $\mathcal{M}$  and  $\mathcal{R}$  the recurrence relation*

$$\mathbb{E}[\tilde{W}_n^s] = \sum_{r=0}^s f_{n,s,r} \mathbb{E}[\tilde{W}_{n-1}^r], \quad n \geq 1,$$

with initial values  $\mathbb{E}[\tilde{W}_0^s] = W_0^s$ , and the values  $f_{n,s,r}$  being given by

$$f_{n,s,r} = \sum_{j=s-r}^s \binom{s}{j} \Lambda^j \sum_{\ell=j-(s-r)}^j \binom{\ell}{j-(s-r)} \left( \frac{a_m(n-1)}{1-\Lambda} \right)^{\ell-j+s-r} \\ \times \sum_{i=\ell}^j \binom{j}{i} \frac{\sigma^i a_m^{j-i}}{m^i (\Lambda-1)^{j-i}} p_{n;(i,\ell)},$$

for  $0 \leq r \leq s$  with

$$p_{n;(i,\ell)} = \begin{cases} \sum_{h=\ell}^{i^*} (-1)^{h-i} \frac{\begin{bmatrix} h \\ \ell \end{bmatrix} \begin{Bmatrix} i \\ h \end{Bmatrix} \binom{m}{h}}{\binom{T_{n-1}}{h}}, & \text{model } \mathcal{M}; \\ \begin{Bmatrix} i \\ \ell \end{Bmatrix} \frac{m^\ell}{T_{n-1}^\ell}, & \text{model } \mathcal{R}. \end{cases}$$

with  $i^* = \min\{i, m\}$ .

**Remark 2.** One readily obtains concrete expressions for  $f_{n,s,r}$  and  $s = 1, 2, \dots$  using the formula above and preferentially a Computer Algebra System. For example, for  $s = 1$  we obtain the model-independent result

$$f_{n,1,0} = -\frac{a_m \Lambda T_0}{(1-\Lambda)T_{n-1}}, \quad f_{n,1,1} = 1 + \frac{\sigma \Lambda}{T_{n-1}}.$$

Note that for  $s \geq 2$  the values  $f_{n,s,r}$  are model-dependent. For triangular urns, the values  $f_{n,s,r}$  can be simplified: One readily observes that  $f_{n,s,0} = 0$ , and for  $1 \leq r \leq s$  we have

$$f_{n,s,r} = \sum_{j=s-r}^s \binom{s}{j} \frac{\Lambda^j \sigma^j}{m^j} p_{n;(j,j-(s-r))}.$$

*Proof.* From (3) and (6) we obtain the distributional equation

$$\tilde{W}_n \stackrel{\mathcal{L}}{=} \tilde{W}_{n-1} + \frac{\sigma \Lambda}{m} \sum_{k=0}^m k \mathbb{I}_n(W^k B^{m-k}) + a_m \frac{\Lambda}{\Lambda-1}. \quad (12)$$

Taking the  $s$ th power in (12) leads to

$$\tilde{W}_n^s \stackrel{\mathcal{L}}{=} \sum_{j=0}^s \binom{s}{j} \tilde{W}_{n-1}^{s-j} \Lambda^j \left[ \sum_{i=0}^j \binom{j}{i} \frac{\sigma^i a_m^{j-i}}{m^i (\Lambda-1)^{j-i}} \sum_{k=0}^m k^i \mathbb{I}_n(W^k B^{m-k}) \right].$$

We take the conditional expectation and simplify the sum

$$\sum_{k=0}^m k^i \mathbb{E}[\mathbb{I}_n(W^k B^{m-k}) \mid \mathbb{F}_{n-1}].$$

By (7) and (8) and the properties of the binomial and hypergeometric distributions we have

$$\sum_{k=0}^m k^i \mathbb{E}[\mathbb{I}_n(W^k B^{m-k}) \mid \mathbb{F}_{n-1}] = \sum_{\ell=0}^i p_{n;(i,\ell)} W_{n-1}^\ell,$$



with  $p_{n;(i,\ell)}$  as given in Lemma 2. Finally, converting  $W_{n-1}^\ell$  into powers of  $\tilde{W}_{n-1}$  and several summation changes lead to the stated result.  $\square$

A direct consequence of Lemma 2 is an expression of the moments  $\mathbb{E}[\tilde{W}_n^s]$  in terms of the moments  $\mathbb{E}[\tilde{W}_\ell^r]$ ,  $1 \leq \ell \leq n-1$  and  $0 \leq r \leq s-1$ .

**Proposition 1.** *The positive integer moments  $\mathbb{E}[\tilde{W}_n^s]$  satisfy the recurrence relation*

$$\mathbb{E}[\tilde{W}_n^s] = \left( \prod_{j=1}^n f_{j,s,s} \right) \left( W_0^s + \sum_{\ell=1}^n \frac{\sum_{i=0}^{s-1} f_{\ell,s,i} \mathbb{E}[\tilde{W}_{\ell-1}^i]}{\prod_{j=1}^{\ell} f_{j,s,s}} \right),$$

with  $f_{n,s,r}$  as given in Lemma 2; in particular

$$f_{n,s,s} = \sum_{j=0}^{s^*} \binom{s}{j} C^j p_{n;(j,j)}, \quad \text{with } C = \frac{\Lambda \sigma}{m},$$

and  $s^* = \min\{s, m\}$ .

*Proof.* Lemma 2 states that

$$\mathbb{E}[\tilde{W}_n^s] = \sum_{r=0}^s f_{n,s,r} \mathbb{E}[\tilde{W}_{n-1}^r] = f_{n,s,s} \mathbb{E}[\tilde{W}_{n-1}^s] + \sum_{r=0}^{s-1} f_{n,s,r} \mathbb{E}[\tilde{W}_{n-1}^r].$$

Consequently, we can write

$$\frac{\mathbb{E}[\tilde{W}_n^s]}{\prod_{j=1}^n f_{j,s,s}} = \frac{\mathbb{E}[\tilde{W}_{n-1}^s]}{\prod_{j=1}^{n-1} f_{j,s,s}} + \sum_{r=0}^{s-1} \frac{f_{n,s,r} \mathbb{E}[\tilde{W}_{n-1}^r]}{\prod_{j=1}^n f_{j,s,s}}.$$

This implies that

$$\frac{\mathbb{E}[\tilde{W}_n^s]}{\prod_{j=1}^n f_{j,s,s}} = \mathbb{E}[\tilde{W}_0^s] + \sum_{\ell=1}^n \sum_{r=0}^{s-1} \frac{f_{\ell,s,r} \mathbb{E}[\tilde{W}_{\ell-1}^r]}{\prod_{j=1}^{\ell} f_{j,s,s}}.$$

Multiplication with  $\prod_{j=1}^n f_{j,s,s}$  leads to the stated result.  $\square$

**3.2. Limits of the shifted moments.** Using Proposition 1, we can derive asymptotic expansions of higher moments of  $\tilde{W}_n$ ; compare with [19], where a special model of small-index urns was treated similarly. This would also allow to strengthen our previous results concerning the central limit theorems for small-index urns,  $\Lambda \leq \frac{1}{2}$ , adding convergence of positive integer moments of the number of white balls to the moments of the normal distribution. We omit the involved computational details. As a first application of Proposition 1, we prove the existence of the limits of the normalized moments of  $\tilde{W}_n = W_n - \frac{a_m}{1-\Lambda} n$ , for large-index urns and also for triangular urn models.

**Theorem 1.** For large urn models with  $\Lambda > \frac{1}{2}$  or triangular urn models with  $a_m = 0$  the moments  $E_s = \lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{\tilde{W}_n^s}{n^{s\Lambda}} \right]$  exist. For  $s = 1$  we obtain

$$E_1 = \left( W_0 - \frac{a_m T_0}{1 - \Lambda} \right) \frac{\Gamma(\frac{T_0}{\sigma})}{\Gamma(\frac{T_0}{\sigma} + \Lambda)}.$$

For the higher moments  $E_s$ ,  $s \geq 1$ , we obtain for model  $\mathcal{M}$  the expressions

$$E_s = \left( \prod_{\ell=1}^{s^*} \frac{\Gamma(\frac{T_0+1-\ell}{\sigma})}{\Gamma(\lambda_{\ell,s})} \right) \left( W_0^s + \sum_{\ell=1}^{\infty} \frac{\sum_{i=0}^{s-1} f_{\ell,s,i} \mathbb{E}[\tilde{W}_{\ell-1}^i]}{\prod_{j=1}^{\ell} f_{j,s,s}} \right),$$

with  $s^* = \min\{s, m\}$ , and  $\lambda_{\ell,s}$  denoting the negated roots of the monic polynomials

$$P_s(x) = \frac{s^*!}{\sigma^{s^*}} \sum_{\ell=0}^{s^*} \frac{\binom{s}{\ell}}{\binom{s^*}{\ell}} C^{\ell} \binom{m}{\ell} \binom{x\sigma + T_0 - \ell}{s^* - \ell}, \quad C = \frac{\Lambda\sigma}{m}.$$

For model  $\mathcal{R}$  we obtain

$$E_s = \left( \frac{\Gamma^s(\frac{T_0}{\sigma})}{\prod_{\ell=1}^s \Gamma(\mu_{\ell,s})} \right) \left( W_0^s + \sum_{\ell=1}^{\infty} \frac{\sum_{i=0}^{s-1} f_{\ell,s,i} \mathbb{E}[\tilde{W}_{\ell-1}^i]}{\prod_{j=1}^{\ell} f_{j,s,s}} \right),$$

with  $\mu_{\ell,s}$  denoting the negated roots of the polynomials

$$Q_s(x) = \frac{1}{\sigma^s} \sum_{\ell=0}^s \binom{s}{\ell} C^{\ell} \binom{m}{\ell} \ell! (x\sigma + T_0)^{s-\ell}.$$

**Remark 3.** The result above implies that for  $s \geq 2$  the limits  $E_s$  of the normalized moments of  $\tilde{W}_n = W_n - \frac{a_m}{1-\Lambda} n$  can be expressed as nested infinite sums. In particular, the second moments are for both models readily obtained as a double sum, additionally using Proposition 1 for  $f_{n,s,r}$  and Lemma 1 for  $\mathbb{E}[\tilde{W}_n]$ . This allows to obtain a closed form expression for the variance of  $\tilde{W}_n/n^{\Lambda}$ , and also its limit  $E_2 - E_1^2 > 0$ . For  $m \geq 2$  it seems that these nested infinite sums cannot be simplified in general due to the structure of the monic polynomials  $P_s(x)$  and  $Q_s(x)$  and their negated roots. In contrast, for  $m = 1$  the roots are much simpler and simplifications occur; compare with the discussion in [6].

*Proof.* In order to prove the existence of  $E_k$ ,  $k \geq 1$  we first turn to the case  $k = 1$ . By Lemma 1 we readily obtain  $\mathbb{E}[\tilde{W}_n]$  and the stated result for  $E_1$ . Next we use Proposition 1 and we proceed in two steps. First, we derive an asymptotic expansion of the products  $\prod_{j=1}^n f_{j,s,s}$  as given in Lemma 2. Then, we inductively prove the existence of the infinite sums in the expressions for  $E_s$ ,  $s > 1$ . Concerning the asymptotic expansions of  $\prod_{j=1}^n f_{j,s,s}$ , we consider for model  $\mathcal{M}$  the relation

$$f_{j,s,s} = \sum_{\ell=0}^{s^*} \binom{s}{\ell} C^{\ell} \frac{\binom{m}{\ell}}{\binom{T_{j-1}}{\ell}} = \frac{\sum_{\ell=0}^{s^*} \frac{\binom{s}{\ell}}{\binom{s^*}{\ell}} C^{\ell} \binom{m}{\ell} \binom{T_{j-1}-\ell}{s^*-\ell}}{\binom{T_{j-1}}{s^*}} = \frac{\sum_{\ell=0}^{s^*} \frac{\binom{s}{\ell}}{\binom{s^*}{\ell}} C^{\ell} \binom{m}{\ell} \binom{(j-1)\sigma + T_0 - \ell}{s^* - \ell}}{\binom{(j-1)\sigma + T_0}{s^*}}.$$

The monic polynomials  $P_s(x)$ , as stated in Theorem 1, are related to  $f_{j,s,s}$  by

$$f_{j,s,s} = \frac{P_s(j-1)}{\prod_{\ell=1}^{s^*} (j-1 + \frac{T_0+1-\ell}{\sigma})}.$$

Let  $\lambda_{\ell,s}$  denote the negated roots of the equation  $P_s(x) = 0$ , such that  $P_s(x) = \prod_{\ell=1}^{s^*} (x + \lambda_{\ell,s})$ . Consequently,

$$f_{j,s,s} = \frac{\prod_{\ell=1}^{s^*} (j-1 + \lambda_{\ell,s})}{\prod_{\ell=1}^{s^*} (j-1 + \frac{T_0+1-\ell}{\sigma})}, \quad \text{and} \quad \prod_{j=1}^n f_{j,s,s} = \prod_{\ell=1}^{s^*} \frac{\Gamma(n + \lambda_{\ell,s}) \Gamma(\frac{T_0+1-\ell}{\sigma})}{\Gamma(\lambda_{\ell,s}) \Gamma(n + \frac{T_0+1-\ell}{\sigma})}.$$

Note that by definition,  $f_{j,s,s} > 0$  for all  $j \in \mathbb{N}$ . Hence, the negated roots  $\lambda_{\ell,s} \neq -n$ , for all  $n \in \mathbb{N}$  and the expression above is well defined. By Stirling approximation, we obtain the asymptotic expansion

$$\prod_{j=1}^n f_{j,s,s} = \frac{n^{\sum_{\ell=1}^{s^*} \lambda_{\ell,s}}}{n^{\frac{s^* T_0 - \binom{s^*}{2}}{\sigma}}} \frac{\prod_{\ell=1}^{s^*} \Gamma(\frac{T_0+1-\ell}{\sigma})}{\Gamma(\lambda_{\ell,s})} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

Let  $[x^k]$  denote the extraction of coefficients operator. Since

$$\begin{aligned} \sum_{\ell=1}^s \lambda_{\ell,s} &= [x^{s^*-1}] \prod_{\ell=1}^{s^*} (x + \lambda_{\ell,s}) = [x^{s^*-1}] \frac{s^*!}{\sigma^{s^*}} \sum_{\ell=0}^{s^*} \frac{\binom{s}{\ell}}{\binom{s^*}{\ell}} C^\ell \binom{m}{\ell} \binom{x\sigma + T_0 - \ell}{s^* - \ell} \\ &= \frac{s^*!}{\sigma^{s^*}} [x^{s^*-1}] \left( \binom{x\sigma + T_0}{s^*} + \frac{s}{s^*} C m \binom{x\sigma + T_0 - 1}{s^* - 1} \right) \\ &= \frac{s^*!}{\sigma^{s^*}} \left( \frac{\sigma^{s^*-1} (s T_0 - \binom{s^*}{2})}{s^*!} + \frac{s C m \sigma^{s^*-1}}{s^* (s^* - 1)!} \right) \\ &= \frac{s^* T_0 - \binom{s^*}{2}}{\sigma} + \frac{C m s}{\sigma} \\ &= \frac{s^* T_0 - \binom{s^*}{2}}{\sigma} + \Lambda s, \end{aligned}$$

it follows that  $n^{\sum_{\ell=1}^s \lambda_{\ell,s}} = n^{\Lambda s} n^{\frac{s^* T_0 - \binom{s^*}{2}}{\sigma}}$ . Hence, we get for model  $\mathcal{M}$  the asymptotic expansion:

$$\prod_{j=1}^n f_{j,s,s} = n^{\Lambda s} \prod_{\ell=1}^{s^*} \frac{\Gamma(\frac{T_0+1-\ell}{\sigma})}{\Gamma(\lambda_{\ell,s})} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

Concerning model  $\mathcal{R}$  we can proceed in a similar fashion. We have

$$f_{j,s,s} = \sum_{\ell=0}^s \binom{s}{\ell} c^\ell \frac{m^\ell}{T_{j-1}^\ell} = \frac{\sum_{\ell=0}^s \binom{s}{\ell} c^\ell \binom{m}{\ell} \ell! T_{j-1}^{s-\ell}}{T_{j-1}^s} = \frac{\sum_{\ell=0}^s \binom{s}{\ell} c^\ell \binom{m}{\ell} \ell! ((j-1)\sigma + T_0)^{s-\ell}}{((j-1)\sigma + T_0)^s},$$

and consider the monic polynomials  $Q_s(x)$  with

$$f_{j,s,s} = \frac{Q_s(j-1)}{(j-1 + \frac{T_0}{\sigma})^s}.$$

Using arguments similar to model  $\mathcal{M}$  we obtain

$$\prod_{j=1}^n f_{j,s,s} = n^{\Lambda s} \frac{\Gamma^s\left(\frac{T_0}{\sigma}\right)}{\prod_{\ell=1}^s \Gamma(\mu_{\ell,s})} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

It remains to prove the existence of the moments  $E_s = \lim_{n \rightarrow \infty} \mathbb{E}\left[\frac{\tilde{W}_n^s}{n^{s\Lambda}}\right]$ . By the asymptotic expansions of  $\prod_{j=1}^n f_{j,s,s}$  the limit  $\lim_{n \rightarrow \infty} \frac{\prod_{j=1}^n f_{j,s,s}}{n^{s\Lambda}}$  exists, and we have

$$\lim_{n \rightarrow \infty} \mathbb{E}\left[\frac{\tilde{W}_n^s}{n^{s\Lambda}}\right] = \lim_{n \rightarrow \infty} \frac{\prod_{j=1}^n f_{j,s,s}}{n^{s\Lambda}} \left(W_0^s + \sum_{\ell=1}^{\infty} \frac{\sum_{i=0}^{s-1} f_{\ell,s,i} \mathbb{E}[\tilde{W}_{\ell-1}^i]}{\prod_{j=1}^{\ell} f_{j,s,s}}\right).$$

We have to show that the sums

$$\sum_{\ell=1}^n \frac{\sum_{i=0}^{s-1} f_{\ell,s,i} \mathbb{E}[\tilde{W}_{\ell-1}^i]}{\prod_{j=1}^{\ell} f_{j,s,s}}$$

are convergent, for  $n \rightarrow \infty$ . Assume inductively that the moments  $\mathbb{E}[\tilde{W}_n^s]$  satisfy  $\mathbb{E}[\tilde{W}_n^s] \sim \kappa_s n^{s\Lambda}$  for some values  $\kappa_s$ . For  $s = 1$ , this readily follows from the asymptotic expansion of  $\mathbb{E}[W_n]$  in Lemma 1. For  $s > 1$  we analyze the recurrence relation stated in Proposition 1:

$$\mathbb{E}[\tilde{W}_n^s] = \left(\prod_{j=1}^n f_{j,s,s}\right) \left(W_0^s + \sum_{\ell=1}^n \frac{\sum_{i=0}^{s-1} f_{\ell,s,i} \mathbb{E}[\tilde{W}_{\ell-1}^i]}{\prod_{j=1}^{\ell} f_{j,s,s}}\right).$$

We already know that there exist constants  $c_s > 0$ , such that  $\prod_{j=1}^n f_{j,s,s} > c_s n^{\Lambda s}$ ,  $s \in \mathbb{N}$ . Since  $p_{n,(i,\ell)} = \mathcal{O}\left(\frac{1}{T_{n-1}^{\ell}}\right) = \mathcal{O}\left(\frac{1}{n^{\ell}}\right)$ , we obtain the crude bound  $f_{n,s,r} = \mathcal{O}(1)$ . By our induction assumption  $\mathbb{E}[\tilde{W}_n^r] \sim \kappa_r n^{r\Lambda}$  for all  $1 \leq r < s$ . We split the sum  $\sum_{i=0}^{s-1} f_{\ell,s,i} \mathbb{E}[\tilde{W}_{\ell-1}^i]$  into two parts:  $\sum_{i=0}^{s-2} f_{\ell,s,i} \mathbb{E}[\tilde{W}_{\ell-1}^i]$  and  $f_{\ell,s,s-1} \mathbb{E}[\tilde{W}_{\ell-1}^{s-1}]$ . Hence, for large-index urn models with  $\Lambda > \frac{1}{2}$ , we obtain for the first part

$$\sum_{\ell=1}^n \frac{\sum_{i=0}^{s-2} f_{\ell,s,i} \mathbb{E}[\tilde{W}_{\ell-1}^i]}{\prod_{j=1}^{\ell} f_{j,s,s}} = \mathcal{O}\left(\sum_{\ell=1}^n \sum_{i=0}^{s-2} \frac{\ell^{i\Lambda}}{\ell^{\Lambda s}}\right) = \mathcal{O}\left(\sum_{\ell=1}^n \frac{1}{\ell^{2\Lambda}}\right) = \mathcal{O}(1).$$

For triangular urns we use Remark 2 and get:

$$f_{n,s,r} = \sum_{j=s-r}^s \binom{s}{j} \frac{\Lambda^j \sigma^j}{m^j} p_{n;(j,j-(s-r))} = \frac{\Lambda^{s-r} \sigma^{s-r}}{m^{s-r}} p_{n;(j,0)} + \mathcal{O}\left(\frac{1}{n}\right) = \mathcal{O}\left(\frac{1}{n}\right),$$

since  $p_{n;(j,0)} = 0$  for  $j > 0$  by definition of the Stirling numbers of the first and second kind.

Concerning the second part  $f_{\ell,s,s-1} \mathbb{E}[\tilde{W}_{\ell-1}^{s-1}]$  we have to show that

$$\sum_{\ell=1}^n \frac{f_{\ell,s,s-1} \mathbb{E}[\tilde{W}_{\ell-1}^{s-1}]}{\prod_{j=1}^{\ell} f_{j,s,s}}$$

exists. For triangular urns we already observed the bound  $f_{n,s,s-1} = \mathcal{O}(\frac{1}{n})$ . For large urns we refine the bound  $f_{n,s,s-1} = \mathcal{O}(1)$  in the following way:

$$\begin{aligned} f_{n,s,s-1} &= \sum_{j=1}^s \binom{s}{j} \Lambda^j \sum_{\ell=j-1}^j \binom{\ell}{j-1} \left( \frac{a_m(n-1)}{1-\Lambda} \right)^{\ell-j+1} \sum_{i=\ell}^j \binom{j}{i} \frac{\sigma^i a_m^{j-i}}{m^i (\Lambda-1)^{j-i}} p_{n;(i,\ell)} \\ &= \binom{s}{1} \Lambda^1 \sum_{\ell=0}^1 \binom{\ell}{0} \left( \frac{a_m(n-1)}{1-\Lambda} \right)^{\ell} \sum_{i=\ell}^1 \binom{1}{i} \frac{\sigma^i a_m^{1-i}}{m^i (\Lambda-1)^{1-i}} p_{n;(i,\ell)} + \mathcal{O}\left(\frac{1}{n}\right). \end{aligned}$$

Since  $p_{n;(j,0)} = 0$  for  $j > 0$  and  $p_{n;(0,0)} = 1$  we get

$$f_{n,s,s-1} = \binom{s}{1} \Lambda^1 \left( \frac{a_m}{\Lambda-1} + \frac{a_m(n-1)}{1-\Lambda} \times \frac{\sigma}{m} \times \frac{m}{T_{n-1}} \right) + \mathcal{O}\left(\frac{1}{n}\right).$$

We have

$$\begin{aligned} \frac{a_m}{\Lambda-1} + \frac{a_m(n-1)}{1-\Lambda} \times \frac{\sigma}{m} \times \frac{m}{T_{n-1}} &= \frac{a_m}{\Lambda-1} + \frac{a_m \sigma (n-1)}{(1-\Lambda)(\sigma(n-1) + T_0)} \\ &= \frac{a_m}{\Lambda-1} + \frac{a_m}{1-\Lambda} \left( 1 + \mathcal{O}\left(\frac{1}{n}\right) \right) \\ &= \mathcal{O}\left(\frac{1}{n}\right), \end{aligned}$$

proving that  $f_{n,s,s-1} = \mathcal{O}(\frac{1}{n})$ . This implies that

$$\sum_{\ell=1}^n \frac{f_{\ell,s,s-1} \mathbb{E}[\tilde{W}_{\ell-1}^{s-1}]}{\prod_{j=1}^{\ell} f_{j,s,s}} = \sum_{\ell=1}^n \mathcal{O}\left(\frac{1}{\ell^{\Lambda+1}}\right) = \mathcal{O}(1).$$

Consequently, the sums

$$\sum_{\ell=1}^n \frac{\sum_{i=0}^{s-1} f_{\ell,s,i} \mathbb{E}[\tilde{W}_{\ell-1}^i]}{\prod_{j=1}^{\ell} f_{j,s,s}}$$

are convergent, for  $n \rightarrow \infty$  and the moments  $\mathbb{E}[\tilde{W}_n^s] \sim \kappa_s n^{s\Lambda}$ .  $\square$

**3.3. Higher moments for large-index urns and triangular urns.** Next we relate the limits  $E_s = \lim_{n \rightarrow \infty} \mathbb{E}\left[\frac{\tilde{W}_n^s}{n^{s\Lambda}}\right]$  of the normalized moments of  $\tilde{W}_n = W_n - \frac{a_m}{1-\Lambda} n$  with the moments of  $\mathcal{W}_\infty$  for large-index urns and with  $\mathfrak{W}_\infty$  for triangular urn models.

**Proposition 2.** *For large-index urns the positive integers moments of  $\mathcal{W}_\infty$  exist and can be expressed in terms of  $E_s = \lim_{n \rightarrow \infty} \mathbb{E}\left[\frac{\tilde{W}_n^s}{n^{s\Lambda}}\right]$ , and*

$$\mathbb{E}[\mathcal{W}_\infty^s] = \sum_{k=0}^s \binom{s}{k} \frac{\Gamma^k\left(\frac{T_0}{\sigma} + \Lambda\right)}{\Gamma^k\left(\frac{T_0}{\sigma}\right)} E_k \left( \frac{a_m T_0}{\sigma(1-\Lambda)} - W_0 \right)^{s-k}.$$

For triangular urn models the positive integers moments of  $\mathfrak{W}_\infty$  exist and can be expressed in terms of  $E_s = \lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{\tilde{W}_n^s}{n^{s\Lambda}} \right]$ :

$$\mathbb{E}[\mathfrak{W}_\infty^s] = \frac{\Gamma^s(\frac{T_0}{\sigma} + \Lambda)}{\Gamma^s(\frac{T_0}{\sigma})} E_s.$$

**Remark 4.** Starting with the second moments, a simple formula for  $\mathbb{E}[\mathcal{W}_\infty^s]$  or  $\mathbb{E}[\mathfrak{W}_\infty^s]$  seems to be elusive, due to the infinite sums in the expressions for  $E_s$  in Theorem 1. We will present different explicit expressions for the moments in the next section.

*Proof.* We have

$$\begin{aligned} \mathcal{W}_n &= g_n(W_n - \mathbb{E}[W_n]) \\ &= g_n n^\Lambda \frac{W_n - n \frac{a_m}{1-\Lambda}}{n^\Lambda} + g_n \left( n \frac{a_m}{1-\Lambda} - \mathbb{E}[W_n] \right) \\ &= g_n n^\Lambda \frac{\tilde{W}_n}{n^\Lambda} - g_n \left( \mathbb{E}[W_n] - n \frac{a_m}{1-\Lambda} \right). \end{aligned} \tag{13}$$

This implies that

$$\mathbb{E}[\mathcal{W}_n^s] = \sum_{k=0}^s \binom{s}{k} (g_n n^\Lambda)^k \mathbb{E} \left[ \frac{\tilde{W}_n^k}{n^{k\Lambda}} \right] \left( g_n \left( n \frac{a_m}{1-\Lambda} - \mathbb{E}[W_n] \right) \right)^{s-k}.$$

We obtain from the asymptotic expansions of  $g_n$  and  $\mathbb{E}[W_n]$  in Lemma 1, and also the previous Lemma the stated result:

$$\lim_{n \rightarrow \infty} g_n n^\Lambda = \frac{\Gamma(\frac{T_0}{\sigma} + \Lambda)}{\Gamma(\frac{T_0}{\sigma})}, \quad \lim_{n \rightarrow \infty} g_n \left( n \frac{a_m}{1-\Lambda} - \mathbb{E}[W_n] \right) = \frac{a_m T_0}{\sigma(1-\Lambda)} - W_0.$$

□

#### 4. AN EXPLICIT EXPRESSION FOR THE MOMENTS

In order to derive explicit expressions for the moments of triangular urns and large-index urns we study first the general solution of certain types of recurrence relations. We use a correspondence between the recurrence relations and paths in weighted directed acyclic graphs in order to obtain explicit solutions using a lattice paths counting argument.

**4.1. Triangular urns.** In the following we determine for triangular urn models an explicit expression for  $\mathbb{E}[\tilde{W}_n^s]$  and also  $E_s = \lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{\tilde{W}_n^s}{n^{s\Lambda}} \right]$ . Note that  $a_m = 0$ , so  $\tilde{W}_n = W_n$ .

**Lemma 3.** Let the sequence  $(e_{n,s})_{n \geq 0, s \geq 1}$  be defined by the recurrence relation

$$e_{n,s} = \sum_{i=1}^s f_{n,s,i} e_{n-1,i},$$

$n \geq 1$ , with initial values  $e_{0,s} = x^s$ ,  $s \in \mathbb{N}_0$ , for a given triple sequence  $(f_{n,s,i})_{n \in \mathbb{N}_0, s \in \mathbb{N}, 1 \leq i \leq s}$ . Then  $e_{n,s} = \sum_{k=1}^s \varphi_{n,s,k} x^k$  is a polynomial in  $x$  of degree  $s$  with no constant term, with  $\varphi_{n,s,k}$  defined by the recurrence relation  $\varphi_{n,s,k} = \sum_{\ell=k}^s f_{n,s,\ell} \varphi_{n-1,\ell,k}$ ,  $n \geq 1$ ,  $s \geq 1$ , and  $\varphi_{0,s,k} = \delta_{s,k}$ .

*Proof.* The statement is by definition true for  $n = 0$  and arbitrary  $s \in \mathbb{N}$ . Assuming the statement for values less than  $n$ , we obtain

$$e_{n,s} = \sum_{k=1}^s \varphi_{n,s,k} x^k = \sum_{i=1}^s f_{n,s,i} e_{n-1,i} = \sum_{i=1}^s f_{n,s,i} \sum_{k=1}^i \varphi_{n-1,i,k} x^k = \sum_{k=0}^s x^k \sum_{i=k}^s f_{n,s,i} \varphi_{n-1,i,k}.$$

By comparison of coefficients of the powers of  $x$ , we obtain the given recurrence relations, which proves the stated result.  $\square$

**Lemma 4.** The coefficients  $\varphi_{n,s,k}$  in the expansion of  $e_{n,s} = \sum_{k=1}^s \varphi_{n,s,k} x^k$  are given by

$$\varphi_{n,s,s} = \prod_{j=1}^n f_{j,s,s}, \quad \varphi_{n,s,s-1} = \sum_{i=1}^n \left( \prod_{j_1=i+1}^n f_{j_1,s,s} \right) f_{i_1,s,s-1} \left( \prod_{j_2=1}^{i_1-1} f_{j_2,s-1,s-1} \right),$$

and in general, for  $1 \leq k \leq s-1$ , by the expression

$$\varphi_{n,s,k} = \sum_{\ell=1}^{s-k} \sum_{\substack{\sum_{\nu=1}^{\ell} h_{\nu}=s-k \\ h_{\nu} \geq 1}} \left[ \sum_{1 \leq i_{\ell} < \dots < i_1 \leq n} \prod_{g=1}^{\ell+1} \left( f_{i_g, s-H_{g-1}, s-H_g} \prod_{j=i_g+1}^{i_{g-1}-1} f_{j, s-H_{g-1}, s-H_g-1} \right) \right];$$

here  $H_k = \sum_{\nu=1}^k h_{\nu}$ , and  $i_0 = n+1$ ,  $i_{\ell+1} = 0$ . We use the convention  $f_{0,s,i} = 1$ .

*Proof.* In order to derive the stated expressions we use a lattice path counting argument. Given a triple sequence  $(f_{n,s,i})_{n \in \mathbb{N}_0, s \in \mathbb{N}, 1 \leq i \leq s}$  we consider a weighted directed acyclic graph  $G = (V, E)$  with vertices  $v \in V$  identified by their pair of labels  $v = (k, \ell)$ , for  $0 \leq k \leq n$  and  $1 \leq \ell \leq s$ . The edges  $e \in E$  are directed from vertices  $(k, \ell)$  to  $(k-1, j)$ , with  $1 \leq k \leq n$ ,  $1 \leq \ell \leq s$  and  $1 \leq j \leq \ell$ , and an edge  $e = ((k, \ell) \rightarrow (k-1, j))$  has weight

$$w(e) = w((k, \ell) \rightarrow (k-1, j)) = f_{k,\ell,j}.$$

The weight of a directed path  $\mathbf{p}$  is defined as the product of the edge weights:

$$w(\mathbf{p}) = \prod_{e \in \mathbf{p}} w(e).$$

We have a one-to-one correspondence between the recurrence relation  $e_{n,s} = \sum_{\ell=1}^s f_{n,s,\ell} e_{n-1,\ell}$  and certain paths in the graph  $G$ : We start at the vertex  $(n, s)$ —the source—and end at vertices  $(0, k)$ ,  $1 \leq k \leq s$ —the sinks. Ending at one of the sinks  $(0, k)$  corresponds to reaching the initial value  $x^k$ . The coefficients  $\varphi_{n,k,s}$  appearing in the expansion  $e_{n,s} = \sum_{k=1}^s \varphi_{n,s,k} x^k$  are given by the sum of weights of certain paths,

$$\varphi_{n,k,s} = \sum_{\text{Path } \mathbf{p}: (n,s) \rightarrow (0,k)} w(\mathbf{p}).$$

In order to obtain the weights we consider refinement of the paths from  $(n, s)$  to  $(0, k)$  taking into account the number of changes of the second coordinate which we call jumps. We can have  $\ell$  jumps, with  $1 \leq \ell \leq s - k$ , and the individual jump heights  $h_\nu$ . The total height  $H_\ell = \sum_{\nu=1}^{\ell} h_\nu$  of the jumps has to equal  $s - k$  under the restriction that  $h_\nu \geq 1$ ,  $1 \leq \nu \leq \ell$ . Moreover, as a further refinement we fix the first coordinates  $\mathbf{i} = (i_1, i_2, \dots, i_\ell)$  of the steps where the jumps occur, with  $1 \leq i_\ell < \dots < i_1 \leq n$ . Given a directed path  $\mathbf{p}$  starting at  $(n, s)$  and ending at  $(0, k)$  with  $\ell$  jumps at steps  $\mathbf{i} = (i_1, i_2, \dots, i_\ell)$  of heights  $h_\nu \geq 1$ , with  $\sum_{\nu=1}^{\ell} h_\nu = s - k$ , the weight of such a path  $\mathbf{p}$  is given by

$$w(\mathbf{p}) = \prod_{g=1}^{\ell+1} \left( f_{i_g, s-H_{g-1}, s-H_g} \prod_{j=i_g+1}^{i_{g-1}-1} f_{j, s-H_{g-1}, s-H_{g-1}} \right).$$

Summing over all paths  $\mathbf{p}$ —taking into account all possible first coordinates  $\mathbf{i} = (i_1, i_2, \dots, i_\ell)$  of the steps where the jumps occur and also the the different heights of the jumps—leads to the stated result.  $\square$

Next we combine both results to obtain an explicit representation of the moments.

**Theorem 2.** *For triangular urn models the positive integer moments  $\mathbb{E}[W_n^s]$  are given by*

$$\mathbb{E}[W_n^s] = \sum_{k=1}^s \varphi_{n,s,k} W_0^k,$$

with  $\varphi_{n,s,k}$  as given in Lemma 4. The limits  $E_s = \lim_{n \rightarrow \infty} \mathbb{E}\left[\frac{W_n}{n^\Lambda}\right]^s$  of the positive integer moments of the normalized random variables  $\tilde{W}_n/n^\Lambda = W_n/n^\Lambda$  can be expressed as

$$E_s = \begin{cases} \prod_{\ell=1}^{s*} \frac{\Gamma(\frac{T_0+1-\ell}{\sigma})}{\Gamma(\lambda_{\ell,s})} \sum_{k=1}^s \tilde{\varphi}_{s,k} W_0^k, & \text{model } \mathcal{M}, \\ \frac{\Gamma(\frac{T_0}{\sigma})^s}{\prod_{\ell=1}^s \Gamma(\mu_{\ell,s})} \sum_{k=1}^s \tilde{\varphi}_{s,k} W_0^k, & \text{model } \mathcal{R}, \end{cases}$$

with  $\tilde{\varphi}_{s,k} = \lim_{n \rightarrow \infty} \frac{\varphi_{n,s,k}}{\varphi_{n,s,s}}$ . Moreover,  $\tilde{\varphi}_{s,s} = 1$  and for  $1 \leq k \leq s - 1$  the values  $\tilde{\varphi}_{s,k}$  are convergent infinite sums:

$$\tilde{\varphi}_{s,k} = \sum_{\ell=1}^{s-k} \sum_{\substack{\sum_{\nu=1}^{\ell} h_\nu = s-k \\ h_\nu \geq 1}} \left[ \sum_{1 \leq i_\ell < \dots < i_1 < \infty} f_{i_1, s-H_1} \frac{\prod_{g=2}^{\ell+1} \left( f_{i_g, s-H_{g-1}, s-H_g} \prod_{j=i_g+1}^{i_{g-1}-1} f_{j, s-H_{g-1}, s-H_{g-1}} \right)}{\prod_{j=1}^{i_1} f_{j, s, s}} \right],$$

with  $f_{n,s,r}$  as given in Lemma 2.

*Proof.* By Lemma 2 and (11) the moments  $\mathbb{E}[W_n^s] = \mathbb{E}[\tilde{W}_n^s]$  satisfy a recurrence relation of type discussed in Lemmas 3 and 4, such that  $e_{n,s} = \mathbb{E}[W_n^s]$ ,  $f_{n,s,i}$  as given in Lemma 2, see also Remark 2, with initial value  $x = W_0$ . This leads to the first part of the stated result. We already



know from Theorem 1 that the  $\lim_{n \rightarrow \infty} \mathbb{E}[(\frac{W_n}{n^\Lambda})^s]$  exists, for  $s \geq 1$ . We also discern from the proof of Theorem 1 that the limit

$$\lim_{n \rightarrow \infty} n^{-\Lambda s} \varphi_{n,s,s} = \lim_{n \rightarrow \infty} n^{-\Lambda s} \prod_{j=1}^n f_{j,s,s}$$

exists and is a quotient of products of Gamma functions, as stated above. We factor out  $\varphi_{n,s,s}$ , such that  $n^{-\Lambda s} \varphi_{n,s,k} = n^{-\Lambda s} \varphi_{n,s,s} \frac{\varphi_{n,s,k}}{\varphi_{n,s,s}}$ , and separate the factor corresponding to  $g = 1$  from the product

$$\prod_{g=1}^{\ell+1} \left( f_{i_g, s-H_{g-1}, s-H_g} \prod_{j=i_g+1}^{i_{g-1}-1} f_{j, s-H_{g-1}, s-H_{g-1}} \right).$$

It follows that the nested infinite sums exist. They all share as a common factor the quotient of products of Gamma functions. This proves the stated form of the common factor and also that  $\tilde{\varphi}_{s,s} = 1$ , leading to the stated result.  $\square$

An alternative more compact representation can be obtained by considering the discrete simplexes

$$\Delta_k^n = \{\mathbf{c} = (c_1, \dots, c_n) : c_i \geq 0, 1 \leq i \leq n, \sum_{i=1}^n c_i = k\}.$$

Given  $\mathbf{c} \in \Delta_{s-k}^n$  we interpret  $c_\ell$  as the changes of the second label passing from a node with first label  $\ell$  to  $\ell - 1$ . Let the edges  $e_\ell = e_\ell(\mathbf{c})$  be defined by

$$e_\ell = ((\ell, s - C_{\ell+1}) \rightarrow (\ell, s - C_\ell)),$$

for  $1 \leq \ell \leq n$  with  $C_j = \sum_{i=j}^n c_i$ . Then, a path  $\mathbf{p} = \mathbf{p}(\mathbf{c})$  from  $(n, s)$  to  $(0, k)$  can be obtained by

$$\mathbf{p} = (e_n, e_{n-1}, \dots, e_1).$$

We extend the definition of the weight function  $w$  to elements  $\mathbf{c} \in \Delta_k^n$  by

$$w(\mathbf{c}) = w(\mathbf{p}(\mathbf{c})) = \prod_{j=1}^n f_{j, s-C_{j+1}, s-C_j}.$$

Consequently, the coefficients  $\varphi_{n,s,k}$  can be alternatively written as

$$\varphi_{n,s,k} = \sum_{\text{Path } \mathbf{p}: (n,s) \rightarrow (0,k)} w(\mathbf{p}) = \sum_{\mathbf{c} \in \Delta_{s-k}^n} w(\mathbf{c}).$$

This implies that

$$e_{n,s} = \sum_{k=1}^s x^k \sum_{\mathbf{c} \in \Delta_{s-k}^n} w(\mathbf{c}) = \sum_{k=1}^s x^k \sum_{\mathbf{c} \in \Delta_{s-k}^n} \prod_{j=1}^n f_{j, s-C_{j+1}, s-C_j}.$$

Concerning the limits  $\tilde{\varphi}_{s,k} = \lim_{n \rightarrow \infty} \frac{\varphi_{n,s,k}}{\varphi_{n,s,s}}$  appearing in  $E_s$  we obtain an we readily obtain the alternative expressions:

$$\tilde{\varphi}_{s,k} = \sum_{\mathbf{c} \in \Delta_{s-k}^\infty} \prod_{j=1}^{\infty} \frac{f_{j,s-C_{j+1},s-C_j}}{f_{j,s,s}}.$$

**4.2. Large-index urns.** In order to extend the explicit results for  $\mathbb{E}[\tilde{W}_n^s]$  to large-index urns we require a direct extension of Lemma 3.

**Lemma 5.** *Let the sequence  $(e_{n,s})_{n \geq 0, s \geq 0}$  be defined by the recurrence relation*

$$e_{n,s} = \sum_{i=0}^s f_{n,s,i} e_{n-1,i}, \quad n \geq 1, s \geq 1,$$

with initial values  $e_{0,s} = x^s$ ,  $s \in \mathbb{N}_0$  and  $e_{n,0} = 1$ ,  $n \in \mathbb{N}_0$ , for a given triple sequence  $(f_{n,s,i})_{n \in \mathbb{N}_0, s \in \mathbb{N}_0, 0 \leq i \leq s}$ . Then  $e_{n,s} = \sum_{k=0}^s \varphi_{n,s,k} x^k$  is a polynomial in the variable  $x$  of degree  $s$  with  $\varphi_{n,s,k}$  defined for  $n \geq 1$ ,  $s \geq 1$  and  $0 \leq k \leq s$  by the recurrence relation  $\varphi_{n,s,k} = \sum_{\ell=k}^s f_{n,s,\ell} \varphi_{n-1,\ell,k}$ , and  $\varphi_{n,0,k} = \delta_{k,0}$ .

The proof of the result above is identical to Lemma 3 and is therefore skipped. Note that by the fact  $e_{n,0} = 1$  and  $\varphi_{n,0,0} = 1$  the constant term  $\varphi_{n,s,0}$  satisfies

$$\varphi_{n,s,0} = \sum_{\ell=0}^s f_{n,s,\ell} \varphi_{n-1,\ell,0} = \sum_{\ell=1}^s f_{n,s,\ell} \varphi_{n-1,\ell,0} + f_{n,s,0}.$$

The explicit expressions for  $\varphi_{n,s,k}$ ,  $1 \leq k \leq s$  and  $s \geq 1$  are identical to Lemma 4. What remains is to obtain an expression for the constant term  $\varphi_{n,s,0}$ .

**Lemma 6.** *Let the sequence  $(e_{n,s})_{n \geq 0, s \geq 1}$  be defined as in Lemma 5. In the expansion  $e_{n,s} = \sum_{k=0}^s \varphi_{n,s,k} x^k$  the coefficients  $\varphi_{n,s,k}$  are for  $1 \leq k \leq s$  as explicitly stated in Lemma 4. For  $k = 0$  the constant term  $\varphi_{n,s,0}$  is given by*

$$\varphi_{n,s,0} = \sum_{k=1}^n \sum_{\ell=1}^s \sum_{r=1}^{s-\ell} \sum_{\substack{\sum_{\nu=1}^r h_\nu = s-\ell \\ h_\nu \geq 1}} \left[ \sum_{k+1 \leq i_r < \dots < i_1 \leq n} \prod_{g=1}^{r+1} \left( f_{i_g, s-H_{g-1}, s-H_g} \prod_{j=i_g+1}^{i_{g-1}-1} f_{j, s-H_{g-1}, s-H_{g-1}} \right) \right];$$

here  $H_j = \sum_{\nu=1}^j h_\nu$ , and  $i_0 = n+1$ ,  $i_{r+1} = k$ , with  $H_{r+1} = s$ .

*Proof.* In order to derive the stated expression we use again lattice path counting argument. Given a triple sequence  $(f_{n,s,i})_{n \in \mathbb{N}_0, s \in \mathbb{N}_0, 0 \leq i \leq s}$  we consider a weighted directed acyclic graph  $G = (V, E)$  with vertices  $v \in V$  identified by their pair of labels  $v = (k, \ell)$ , for  $0 \leq k \leq n$  and  $0 \leq \ell \leq s$ . The edges  $e \in E$  are directed from vertices  $(k, \ell)$  to  $(k-1, j)$ , with  $1 \leq k \leq n$ ,  $1 \leq \ell \leq s$  and  $0 \leq j \leq \ell$ . The main difference to the triangular urn models is the appearance of edges  $(k, \ell)$  to  $(k-1, 0)$ ,  $1 \leq k \leq n$  and  $1 \leq \ell \leq s$ , which will contribute to constant term  $\varphi_{n,s,0}$ . We have a one-to-one correspondence between the recurrence relation  $e_{n,s} = \sum_{\ell=0}^s f_{n,s,\ell} e_{n-1,\ell}$  and certain paths in the graph  $G$ : We start at the vertex  $(n, s)$ —the source—and end at vertices  $(0, \ell)$ ,

$1 \leq \ell \leq s$ , or  $(k, 0)$ ,  $0 \leq k \leq n-1$ —the sinks. Ending at one of the sinks  $(0, k)$  corresponds to reaching the initial value  $x^k$  and ending at  $(k, 0)$  with  $0 \leq k \leq n-1$  corresponds to reaching the initial value  $1 = x^0$  and contributes to the constant term. The coefficients  $\varphi_{n,k,s}$  are for  $1 \leq k \leq s$  identical to the triangular urn models

$$\varphi_{n,k,s} = \sum_{\text{Path } \mathbf{p}: (n,s) \rightarrow (0,k)} w(\mathbf{p}).$$

The coefficient  $\varphi_{n,0,s}$  is given by the sum of weights of all path ending at one of the sinks  $(k, 0)$ ,  $0 \leq k \leq n-1$ :

$$\varphi_{n,0,s} = \sum_{k=0}^{n-1} \sum_{\text{Path } \mathbf{p}: (n,s) \rightarrow (k,0)} w(\mathbf{p}).$$

The weight of such path are obtained by following a path  $\mathbf{p}$  from  $(n, s)$  to an arbitrary vertex  $(k, \ell)$ , and then end at a sink  $(k-1, 0)$  via the edge weighted by  $f_{k,\ell,0}$ ,  $1 \leq k \leq n$  and  $1 \leq \ell \leq s$ :

$$\varphi_{n,s,0} = \sum_{k=1}^n \sum_{\ell=1}^s f_{k,\ell,0} \sum_{\text{Path } \mathbf{p}: (n,s) \rightarrow (k,\ell)} w(\mathbf{p}).$$

The weight of such path can be calculated as for  $\varphi_{n,s,k}$  by taking into account the position (first coordinate) of the jumps at  $k+1 \leq i_r < \dots < i_1 \leq n$ , and their heights  $h_1, \dots, h_r$ . The weight of such a directed path  $\mathbf{p}$  times the weight of the last step  $f_{k,\ell,0}$  is given by

$$f_{k,\ell,0} \cdot w(\mathbf{p}) = \prod_{g=1}^{r+1} \left( f_{i_g, s-H_{g-1}, s-H_g} \prod_{j=i_g+1}^{i_{g-1}-1} f_{j, s-H_{g-1}, s-H_{g-1}} \right),$$

where we have used the convention  $i_0 = n+1$ ,  $i_{r+1} = k$ , with  $H_{r+1} = s$ . Taking into account all such path leads to the stated result.  $\square$

**Theorem 3.** *For large-index urns the positive integer moments  $\mathbb{E}[\tilde{W}_n^s]$  of the shifted random variable  $\tilde{W}_n = W_n - \frac{na_m}{1-\Lambda}$  are given by*

$$\mathbb{E}[\tilde{W}_n^s] = \sum_{k=0}^s \varphi_{n,s,k} W_0^k,$$

with  $\varphi_{n,s,k}$  for  $1 \leq k \leq s$  as given in Lemma 4 and  $\varphi_{n,s,0}$  as given in Lemma 6. The limits  $E_s = \lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{\tilde{W}_n}{n^\Lambda} \right]^s$  can be expressed as

$$E_s = \begin{cases} \prod_{\ell=1}^{s^*} \frac{\Gamma(\frac{T_0+1-\ell}{\sigma})}{\Gamma(\lambda_{\ell,s})} \sum_{k=0}^s \tilde{\varphi}_{s,k} W_0^k, & \text{model } \mathcal{M}, \\ \frac{\Gamma(\frac{T_0}{\sigma})^s}{\prod_{\ell=1}^s \Gamma(\mu_{\ell,s})} \sum_{k=0}^s \tilde{\varphi}_{s,k} W_0^k, & \text{model } \mathcal{R}, \end{cases}$$

with  $\tilde{\varphi}_{s,k}$  for  $1 \leq k \leq s$  as stated in Theorem 2. Moreover,  $\tilde{\varphi}_{s,0}$  is a convergent infinite sum:

$$\tilde{\varphi}_{s,0} = \sum_{k=1}^{\infty} \sum_{\ell=1}^s \sum_{r=1}^{s-\ell} \sum_{\substack{\sum_{\nu=1}^r h_{\nu}=s-\ell \\ h_{\nu} \geq 1}} \left[ \sum_{k+1 \leq i_r < \dots < i_1 < \infty} \frac{\prod_{g=2}^{r+1} \left( f_{i_g, s-H_{g-1}, s-H_g} \prod_{j=i_g+1}^{i_{g-1}-1} f_{j, s-H_{g-1}, s-H_g-1} \right)}{\prod_{j=1}^{i_1} f_{j, s, s}} \right];$$

## 5. CONCLUSION AND OUTLOOK

**5.1. Summary.** We studied for a part I and II two-color affine linear urn models with multiple drawings—sample size  $m \geq 1$ —under two sampling models the distribution of the number of white balls  $W_n$  after  $n$  draws. Concerning the distribution of the number of white balls  $W_n$  we obtained several limit laws summarized in the Theorem below.

**Theorem 4** (Limit laws for affine balanced two-color urn models). *For a balanced two-color affine urn model with sample size  $m \geq 1$ , let  $\Lambda$  denote the ratio of the two eigenvalues of  $M$ .*

- (1) *Small-index urns, the case  $\Lambda \leq \frac{1}{2}$ :  $\frac{W_n - \mathbb{E}(W_n)}{\sqrt{\mathbb{V}(W_n)}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ .*
- (2) *Large-index urns, the case  $\Lambda > \frac{1}{2}$ :  $\frac{W_n - \mathbb{E}(W_n)}{n^{\Lambda}} \xrightarrow{(a.s.)} L$ .*
- (3) *Triangular urns, the case  $b_0 \cdot a_1 = 0$ :  $\frac{W_n}{n^{\Lambda}} \xrightarrow{(a.s.)} T$ .*

*For large-index and triangular urn models we have convergence of all positive integer moments.*

The central limit theorem for small-index urn is obtained in part I of Theorem 3. The random variables  $L$  and  $T$  are variants of the almost-sure limits  $\mathcal{W}_{\infty}$  and  $\mathfrak{W}_{\infty}$  and their almost-sure convergence can be easily deduced from part I of Theorem 2 and Proposition 4. The convergence of the moments follows directly from Proposition 2 and Theorems 2 and 3.

**5.2. Open problems and extensions.** A natural question is to extend the results to nonaffine linear urn models with multiple drawings, both balanced and unbalanced, and to extend the general limit theorems for  $m = 1$  of Janson [13, 14]. Moreover, it is also of interest to extend the analytic combinatorial framework developed for sample size  $m = 1$  by Flajolet et al. [9, 10] and Morcrette [24] to urn models with multiple drawings, both balanced and unbalanced to obtain precise information about  $W_n$  for fixed  $n$ . Another important question is the study of the limit law of triangular urn models and large-index urns for  $m > 1$  similar to the case  $m = 1$ , which is now well understood [3, 4, 17] in terms of fixed-point equations (smoothing transforms). Does there exist a simpler expression for the moments?

The methods applied in this work do not seem to be easily adapted to unbalanced urn models. See for example [1] for a first step towards the analysis of such models. However, extensions to more than two colors and to certain classes of nonaffine models can be obtained. For example, one can readily generalize the affine linearity condition (2) and the martingale structure of Proposition 1 to balanced urn models with  $r \geq 2$  colors. Moreover, using the correspondence between recurrence relations for integer moments and weighted directed acyclic graphs it seems possible to study more general classes of two-color balanced urn models. It is possible to extend

the approach of Flajolet et al. [9] for balanced urns and of Morcrette [24] for unbalanced urns to derive partial differential equations for suitably defined generating functions; we comment on this elsewhere [21].

## REFERENCES

- [1] R. AGUECH AND O. SELMI (2014). A Generalized Urn Model With Multiple Drawing And Random Addition (I), poster presented at the AofA 2014 - 25th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms Paris, France, June 1620.
- [2] A. BAGCHI AND A. K. PAL (1985). Asymptotic normality in the generalized Pólya-Eggenberger urn model, with an application to computer data structures. *SIAM J. Algebraic Discrete Math.* **6**, 394–405.
- [3] B. CHAUVIN, N. POUYANNE AND R. SAHNOUN (2011). Limit distributions for large Pólya urns. *The Annals of Applied Probability*, **21**, 1–32.
- [4] B. CHAUVIN, N. POUYANNE, AND C. MAILLER (2014). Smoothing equations for large Pólya urns. *Journal of Theoretical Probability* (to appear).
- [5] M.-R. CHEN AND C.-Z. WEI (2005). A New Urn Model, *Journal of Applied Probability* **42**, 964–976, 2005.
- [6] M.-R. CHEN AND M. KUBA (2013). On generalized Polya urn models. *Journal of Applied Probability*, Volume 50, Number 4, 909–1216.
- [7] F. EGGENBERGER AND G. PÓLYA (1923). Über die Statistik verketteter Vorgänge. *Z. Angewandte Math. Mech.* **1**, 279–289.
- [8] P. EHRENFEST AND T. EHRENFEST (1907). Über zwei bekannte Einwände gegen das Boltzmannsche H-theorem. *Physikalische Zeitschrift*, **8**, 311–314.
- [9] P. FLAJOLET, P. DUMAS AND V. PUYHAUBERT (2006). Some exactly solvable models of urn process theory, *Discrete Mathematics and Theoretical Computer Science*, vol. AG, 59–118, 2006, in “Proceedings of Fourth Colloquium on Mathematics and Computer Science”, P. Chassaing Editor.
- [10] P. FLAJOLET, J. GABARRÓ AND H. PEKARI (2005). Analytic urns, *Annals of Probability* **33**, 1200–1233.
- [11] R. L. GRAHAM, D. E. KNUTH, AND O. PATASHNIK, *Concrete Mathematics*, Addison-Wesley, 1994.
- [12] P. HALL AND C. HEYDE (1980). *Martingale Limit Theory and Its Applications*. Academic Press, New York.
- [13] S. JANSON (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic processes and applications*, **110**, 177–245.
- [14] S. JANSON (2006). Limit theorems for triangular urn schemes, *Probability Theory and Related Fields* **134**, 417–452.
- [15] N. L. JOHNSON AND S. KOTZ (1977). *Urn Models and Their Application*. John Wiley, New York.
- [16] N. L. JOHNSON, S. KOTZ, AND H. MAHMOUD (2004). Pólya-type urn models with multiple drawings. *Journal of the Iranian Statistical Society*, **3**, 165–173.
- [17] M. KNAPE AND R. NEININGER (2013+). Pólya urns via the contraction method. *Combinatorics, Probability and Computing (Special issue dedicated to the memory of Philippe Flajolet)* (to appear).
- [18] S. KOTZ AND N. BALAKRISHNAN (1997). Advances in urn models during the past two decades. *Advances in Combinatorial Methods and Applications to Probability and Statistics*, Birkhäuser, Boston, MA, pp. 203–257.
- [19] M. KUBA, H. MAHMOUD AND A. PANHOLZER (2013). Analysis of a generalized Friedman’s urn with multiple drawings. *Discrete Applied Mathematics*, Volume 161, Issue 18, 2968–2984.
- [20] M. KUBA AND H. MAHMOUD (2014+). On urn models with multiple drawings I: urns with a small index. *Submitted*.
- [21] M. KUBA AND B. MORCRETTE (2014+). Analytic combinatorics of urn models with multiple drawings. *Preprint*.
- [22] H. MAHMOUD (2008). *Pólya Urn Models*, Chapman-Hall, Orlando..
- [23] H. MAHMOUD (2013). Drawing multisets of balls from tenable balanced linear urns. *Probability in the Engineering and Informational Sciences*, **27**, 147–162.

- [24] B. MORCRETTE, *Analytic combinatorics and urn models*, Ph.D. thesis, 2013.
- [25] J. MOLER, F. PLO AND H. URMENETA (2013). A generalized Pólya urn and limit laws for the number of outputs in a family of random circuits. *TEST*, **22**, 46–61.
- [26] N. POUYANNE (2008). An algebraic approach to Pólya processes. *Annales de l’Institut Henri Poincaré*, Vol. 44, No. 2, 293–323.
- [27] H. RENLUND (2010). Generalized Pólya urns via stochastic approximation, online available on the arXiv, <http://arxiv.org/abs/1002.3716>.
- [28] R. STANLEY (1997). *Enumerative Combinatorics Volume I*, Cambridge University Press.
- [29] T. TSUKIJI AND H. MAHMOUD (2001). A limit law for outputs in random circuits, *Algorithmica*, 403–412.
- [30] D. WILLIAMS (1991). *Probability with Martingales*. Cambridge University Press, Cambridge, UK.

MARKUS KUBA, INSTITUTE OF APPLIED MATHEMATICS AND NATURAL SCIENCES, UNIVERSITY OF APPLIED SCIENCES-TECHNIKUM WIEN, HÖCHSTÄDTPLATZ 5, 1200 WIEN

*E-mail address:* kuba@technikum-wien.at

HOSAM M. MAHMOUD, DEPARTMENT OF STATISTICS, THE GEORGE WASHINGTON UNIVERSITY, WASHINGTON, D.C. 20052, U.S.A.

*E-mail address:* hosam@gwu.edu